NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

THESIS

ESTIMATING THE NUMBER OF AVAILABLE HIGH QUALITY RECRUITS AT A COUNTY LEVEL

by

Tim Duening

March, 1997

Thesis Advisor: Mike Cook
Associate Advisor: Stephen L. Mehay

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE March 1997	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE Estimating the Number of Available High Quality Recruits at a County Level

5. FUNDING NUMBERS

6. AUTHOR Timothy J. Duening

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

Whether or not a person is available to be recruited is essentially determined by two factors. First, the person has to be desirable to the military in terms of meeting the entry screens. Desirable, as defined by the military, is a person of "high quality." The "high quality" market is defined as high school graduates scoring above the 50th percentile on the Armed Forces Qualification Test (AFQT). The second factor is determined by the individual's choice to attend college. A person who attends college is, for all practical purposes, not included in the military enlistment market. The two factors affecting availability are not independent of each other. A person who scores high on the AFQT is more likely to attend college and therefore be exempt from the potential recruitment pool. This simultaneity must be accounted for in determining the probability that a person is not only qualified but also available for recruitment.

This thesis takes into account the simultaneity of being "high quality" and a non-college attendee in a model that uses alternative demographic and economic demographic variables. These variables include parent's education, family income, single parent household, race and gender. The general findings are that individuals with very low or very high values of parent's education and family income have a lower probability of being in the recruiting pool, whereas those with average values of these characteristics have a higher probability of being in the recruiting pool. This study also finds that minorities were less likely to be in the recruiting pool compared to whites.

14. SUBJECT TERMS Estimation Recruiting County

15. NUMBER OF PAGES 62

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

Approved for public release; distribution is unlimited.

ESTIMATING THE NUMBER OF AVAILABLE HIGH QUALITY RECRUITS AT A COUNTY LEVEL

Timothy J. Duening
Lieutenant Commander, United States Navy
B.S., University of Wisconsin-Stevens Point, 1984

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
March 1997

ABSTRACT

Whether or not a person is available to be recruited is essentially determined by two factors. First, the person has to be desirable to the military in terms of meeting the entry screens. Desirable, as defined by the military, is a person of "high quality." The "high quality" market is defined as high school graduates scoring above the 50th percentile on the Armed Forces Qualification Test (AFQT). The second factor is determined by the individual's choice to attend college. A person who attends college is, for all practical purposes, not included in the military enlistment market. The two factors affecting availability are not independent of each other. A person who scores high on the AFQT is more likely to attend college and therefore be exempt from the potential recruitment pool. This simultaneity must be accounted for in determining the probability that a person is not only qualified but also available for recruitment.

This thesis takes into account the simultaneity of being "high quality" and a non-college attendee in a model that uses alternative demographic and economic explanatory variables. These variables include parent's education, family income, single parent household, race and gender. The general findings are that individuals with very low or very high values of parent's education and family income have a lower probability of being in the recruiting pool, whereas those with average values of these characteristics have a higher probability of being in the recruiting pool. This study also finds that minorities were less likely to be in the recruiting pool compared to whites.

TABLE OF CONTENTS

I. INTRODUCTION

A. BACKGROUND

On 1 July 1973, six months after the final draft call, the military moved from conscription to a total voluntary military service. Since this historic day, much research has been undertaken in the area of recruiting for the military service. The challenge for recruiters has been twofold: (1) to recruit at sufficient levels to maintain the all volunteer force; and (2) to recruit personnel of sufficient quality. Taken separately, either of these challenges can be met with success. However, to meet both challenges simultaneously is an extremely difficult and complex assignment.

Each year, recruiting resources are allocated to such things as advertising, promotional activities and provisions for the support of field recruiting forces. By far, the largest expenditures are on the latter of the three. Enhancing the productivity of recruiters and expenditures on them would have a substantial impact on DoD and the individual services as well. Along with the savings in dedicated resources, the provision of adequate numbers of qualified service recruits is an important determinant of overall defense readiness. While it is important to have information on the qualified recruiting pool nationwide, the estimation of the number of recruits of sufficient quality in specific geographical areas also would be a valuable tool to enhance the productivity of local recruiting districts.

Because not all recruiting districts enjoy a large supply of available "high quality" individuals for potential enlistment, it would be useful to identify areas of high recruiting market potential and areas of low recruiting market potential. An area where test scores are relatively low, or where college attendance rates are relatively high will have a smaller recruiting market potential than an area with of high academic achievement and low college attendance rates. Upon determining the distribution of the potential recruiting pool across local areas, recruiting resources could be more efficiently allocated geographically. The military would not want, for example, to allocate more resources to an area where the characteristics lend themselves to producing a relatively low market potential compared to areas with characteristics conducive to a high percentage of available "high quality" recruits.

B. PROBLEM

Whether or not a person is available to be recruited is essentially determined by two factors. First, the person has to be desirable to the military in terms of meeting the entry screens. Desirable, as defined by the military, is a person of "high quality." The "high quality" market is defined as high school graduates scoring above the 50th percentile on the Armed Forces Qualification Test (AFQT). The second factor is determined by the individual's choice to attend college. A person who attends college is, for all practical purposes, not included in the military enlistment market. The two factors affecting availability, however, are not independent of each other. A person who scores high on the AFQT is more likely to attend college and therefore be exempt from the potential recruitment pool. This simultaneity must be accounted for in determining the probability that a person is not only qualified but also available for recruitment.

Past research has focused solely on "high quality" (as measured by AFQT scores) or on attitudes toward the military (using the Youth Attitude Tracking Surveys) without regard to the simultaneity associated with the decision to attend college or not. A model which estimates the characteristics of a "high quality" recruit is not sufficient to estimate the size of the recruiting pool as it would ignore the interdependence between "high quality" background and college attendance.

This thesis will take into account the simultaneity between of "high quality" and non-college attendance in a model using alternative demographic and economic variables that can be used to estimate the number of available "high quality" recruits in a given geographic area. The geographic area, for the purposes of this study, will be defined as the county.

C. SCOPE OF RESEARCH

This research will use data obtained from the National Educational Longitudinal Survey of 1988 (NELS 88). The NELS 88 surveyed eight grade students in 1988 and conducted three follow-up surveys as of the date of this thesis. The follow-up surveys were conducted in 1990, 1992, and 1994. Using the 1992 second follow-up survey, a cross-section of 1992 high school seniors will be constructed to estimate test scores. Also, using the 1994

third follow-up survey college status (attend or non-attend), of these students can be determined.

This study will essentially be broken down into two stages. In stage I, a probability model will be developed to obtain parameter estimates (β's) of the demographic and economic characteristics of an individual who is "high quality" *and* is not a college attendee. These individual's (high quality/non-college attendee) will make up what will be known as the "target group." The "target group" is what the military both desires to recruit (because they are "high quality"), and who are available for recruitment (not in college). To demonstrate the use of the parameter estimates obtained in stage I, stage II will use these estimates and the associated characteristics of individuals in specific counties to estimate the size of the "target group" available for recruitment in each of these specific counties. County level characteristics are derived from the Public Use Micro-level Survey (PUMS, 5% sample) data set, which is based on the 1990 Decennial Census of the United States.

D. ORGANIZATION OF STUDY

Chapter II will consist of a literature review and will review previous studies that focused on estimating the size of a recruiting pool in a given geographic area. It will also review previous studies that modeled characteristics of individuals and their relationship to aptitude test scores and review studies that modeled characteristics of individuals and their relationship to college attendance. Chapter III will be divided into two major sections. The first section will be a discussion of the two data sets that were used in conducting this research. The second section of Chapter III will discuss the methodology used for model development, and discuss how the estimates obtained in the model can be used to estimate the number of "high quality" individuals available for recruitment at the county level.

Chapter IV will discuss the results of the research and will have a dedicated section for each model. Also included in Chapter IV will be an illustration of how the estimates obtained by model can be used to estimate the number of available high quality individuals at the county level. Chapter V will offer conclusions and recommendations for further study.

II. LITERATURE REVIEW

A. OVERVIEW

This study provides a method which could be used to predict the number of available "high quality" recruits at the county level. This entails developing a single model which predicts the probability of an individual being "high quality" and not in college. The dependent variable used in the probability model developed in this study (high quality/non-college attendee), therefore, is based on two areas of previous research. The first area of research has focused on determinants of an individual's academic achievement, and the second area of research has focused on determinants of individuals' educational attainment. Numerous studies have been conducted in both of these areas and a literature review of both will provide the basis for developing the model used in this study. First, a discussion of literature review of previous studies that estimated the number of available recruits in a given area will be discussed.

B. PREVIOUS STUDIES

An estimation of the number of available recruits in a geographic area that uses simply the number of available youth between the ages of 17 and 21 would certainly suffer from bias. This simple demographic estimation technique would not take into account the number of 17-21 year old's who were of insufficient quality, as not every 17-21 year old would meet the armed forces criteria of being a high school graduate scoring in the top 50 percent on the AFQT. Using only raw estimates of 17-21 year olds involves a second bias in that there is no deduction for those in this age group who attend college, and are therefore unavailable for enlistment. The U.S. Army Recruiting Command has used this simple technique in previous attempts to estimate the number of available recruits in a given geographical area (Thomas and Gorman, 1991).

Moral and Medical qualification rates have also been used to estimate the number of available recruits in a given geographic area. For example, an Air Force study (1985), used national delinquency rates to estimate the number of available recruits in a given geographical area.

Studies by Bock and Moore (1984), Behrendt *et al.* (1986), Curtis, Borack, and Wax (1987) and Orvis and Gahart (1989) have studied the relationship between social and economic characteristics and AFQT test scores. These studies may accurately predict the number of "high quality" individuals in a given geographic area; however, these estimates will not accurately predict the size of a recruiting pool for a given geographic area. Using only the number of "high quality" individuals in a given area to predict the size of a recruiting pool would be an overestimate because this estimate of "high quality" is not the same as an estimate of "high quality" individuals available for recruitment in a given geographic area. Although "high quality" in these studies is accounted for, the simultaneity of being of "high quality" and attending college is not accounted for. These studies are sure to produce an overestimation of available recruits in a given area, as there is no account for the fact that the same characteristics which determine "high quality" also have a similar effect on whether or not a person attends college. If college attendance is not accounted for in an estimation of available recruits, an overestimation will result because those who attend college are unavailable for recruitment. Additionally, areas with a high predicted test scores will also be areas with high rates of college attendance. The results will be relatively more biased for these areas.

A study conducted by Thomas and Gorman (1991), accounted for both high quality and individual interest in joining the military. Although this study did take into account the fact that not all available people between the ages of 17 and 21 would be of high quality, it did not account for the simultaneity associated with being "high quality" and attending college. In their study "Estimation of High-Quality Military Available and Interested" Thomas and Gorman develop a method for estimating high-quality military available (HQ QMA) accounting for the geographic variability in the mental, moral, and physical qualifications required for entry into the military. Further, Thomas and Gorman develop a model to account for both people's interest in joining the military and their subsequent decision to join the military (HQ QMJ). Each of these models will be described separately.

HQ QMA Estimation. The equation derived by Thomas and Gorman is given by:

HQ QMA = (W&P estimate of civilian high school graduates) X
(proportion morally qualified) X
(proportion medically qualified) X
(proportion in categories I-IIIA| county characteristics).

W&P estimate of civilian high school graduates was provided for each county by a private firm; Woods and Poole Economics, Inc. This is an estimation of military available who were not in high school. Twelve estimates for each county were used; white, black, and Hispanic men and women in two age group categories; 17-21 and 22-29 years old.

The proportion of morally qualified was estimated using the moral qualification rates from a U.S. Air Force Personnel Composition Study (1985), which estimated national delinquency rates by gender and estimated the rates as 95.2% for men and 98.4% for women. The proportion of medically qualified was obtained by Laurence (1988) using data from the National Health and Nutrition Examination Survey, 1976-80 (NHANES II), conducted by the National Center for Health Statistics. The moral and medical qualification rates were assumed constant across counties.

The proportion in categories I-IIIA|county characteristics was estimated using a regression equation using various county characteristics to determine the fraction of available population likely to score in the I-IIIA mental categories. The set of explanatory variables used in the regression equation that estimated the fraction of population in categories I-IIIA were parent's education, age of person taking the test, and total net family income. The data used to estimate the AFQT equations was obtained from the 1979 wave of the NLSY.

HQ QMJ Estimation. Another multiplicative model was used to estimate the number of high quality personal who were likely to join the military service. This model (high quality qualified personal likely to join) was derived by using the results from the first model and multiplying them first by a probability of the number of persons interested in joining the military given a mental category group (I-IIIA), then by a probability of the number of persons who would join given a response of interested or not interested on the national YATS survey.

The model to estimate the probability of a person being interested in joining the military given a particular mental category group was developed using a four choice ordered logit model with the level of interest in the military as the ordinal dependent variable. Explanatory variables were mental group, age, parents' education, and poverty level.

The model to estimate the probability of a person joining the military given the persons' stated interest was developed through the use of a binomial logistic regression. The explanatory variables used were age, poverty status, race, AFQT score in categories I-IIIA as well as stated interest responses to a NLSY survey question "Do you think, in the future, that you will..." with the four possible responses of "definitely try to enlist," "probably try to enlist," "probably not try to enlist," and "definitely not try to enlist."

The equation developed by Thomas and Gorman is given by:

$$QMJ = QMA^{m}_{i} \text{ X } Pr(Interest^{m}_{j} \mid Mental\ Group_{i}) \text{ X } Pr(Join^{m} \mid Interest)$$

where:

m = white plus other, black or Hispanic;

i = mental category I&II, mental category IIIA; and

j = definitely interested, probably interested, probably not interested, and definitely not interested.

Although this model is a step in the right direction for accounting for the simultaneity bias which exists between "high quality" and college attendance, there are some drawbacks. First, the measurement of 'interest' is not completely accurate. Not all of those individuals who answered "definitely will" or "probably will" enlist actually enlisted, furthermore, many of the individuals who answered "probably will not enlist" or "definitely will not enlist" actually enlisted.

A second area of bias in this model is that the multiplicative model relies on the assumption that "high quality" and "interested" are independent outcomes. Specifically, Pr (High Quality and Interested) = Pr (High Quality) X Pr(Interested) *only* if "high quality" and "interested" are statistically independent of each other. If a major component of "interested" is the decision not to attend college then this assumption is clearly violated, as will be seen in the results of this thesis. Further, there will be a systematic bias in the

geographic estimates of the recruiting pool. Estimates presented later in this thesis suggest that it is the distribution of characteristics within a county that determine the size of the "target group," not the mean levels of these characteristics.

C. PREVIOUS STUDIES OF ACADEMIC ACHIEVEMENT AND EDUCATIONAL ATTAINMENT

1. Overview

This thesis describes a method which could be used to estimate the number of available recruits in a geographic location, specifically at the county level. The foundation of the model which is developed in this thesis to estimate this number uses a combination of two sets of previous research. The first set of previous research models the relationship between various demographic and economic variables and academic achievement, and the second set of previous research models the relationship between various demographic and economic variables and educational attainment. The dependent variable used in this study to obtain parameter estimates of the various demographic and economic characteristics of a person who is "high quality" and a non-college attendee is a binary variable which equals one if a person is "high quality" *and* not in college, and zero otherwise. Numerous previous studies have been conducted which explore the determinants of academic achievement, while others have explored the determinants of educational attainment. Both will be discussed separately.

2. Previous Studies of Academic Achievement

The relationship between family and demographic characteristics and student test scores is well documented (Koretz, 1987, 1992; Fuchs and Reklis, 1992). There have been several social and demographic characteristics mentioned in the literature that are considered to have adverse effects on family environment and, in turn, student test scores (Zill and Rogers, 1988; Fuchs and Reklis, 1992; Zill, 1992). The most consistent findings of this research are:

1. Educational attainment of parents is strongly related to test scores of their children.

Higher educational attainment of parents has been linked to the provision of a more stimulating home environment and to values that encourage self-direction in a child (Kohn, 1983; Bradley, 1985). This result is generally confirmed in the empirical research. For

example, Hill and O'Neill (1993) include mother's education in their study on test scores and find that an additional year of mother's education raises the average test score by 1.2 percentage points.

2. Family income is positively related to a student's test score.

The relationship between income and achievement is also well documented. Hill and O'Neill (1993) find that income has a positive and significant effect on children's test scores. Specifically, they found that an increase of $10,000 per year would increase test scores by 2.4 percentage points. Hanushek (1992) also finds that permanent income has a systematic effect on achievement.

In 1986, 20 percent of children under the age of 18 were below the poverty line. For black and Hispanic children these rates were 43 and 37 percent, respectively (Simmons, Finlay, and Yang, 1991). Poverty generates several adverse effects on adolescents. These include a 50 percent increase in the likelihood of having physical or mental problems, an increase in the likelihood of becoming victims of violence as well as a sharp increase in high school dropout rate (Simmons, Finlay, and Yang, 1991).

3. The evidence on the effect of living in a single-parent household is mixed.

Hetherington, Camera, and Featherman (1981), find that there are differences favoring children from two parent families in both achievement and grade point average. They note, however, that the differences in achievement are too small to be meaningful. Milne et al. (1986) find that the negative effects on achievement of children living in single parent families are due to variables closely related to single parent families such as income, mothers' employment, parental expectations and parental help with homework.

Krein and Beller (1988), and Cook (1995), found that residing in a single-parent household had a negative effect on student's test scores. Specifically, Krein and Beller find that "the negative effect of living in a single-parent family increases with the number of years spent living in this type of family, is greatest during the preschool years, and is larger for boys than girls."

In contrast, Hanushek (1992) finds there is no effect on achievement of a child of a single parent family once income is controlled for, while Hill and O'Neill (1993) find that

marital status variables have only weak and statistically insignificant effects on children's test scores when factors such as mother's characteristics and family income are accounted for. Desai, Chase-Lansdale, and Michael (1989) also find family structure has little effect on child achievement holding other variables constant.

3. Previous Studies of Educational Attainment

Studies on the determinants of children's educational attainments are numerous. The general results are that a child's educational attainment is strongly related to several family characteristics including family income, parent's educational attainment and family structure (single- versus two-parent families). A recent review "The Determinants of Children's Attainments: A Review of Methods and Findings" (Haveman and Wolfe, 1995) provides a comprehensive literature review of many of the previous studies which developed a model of the characteristics of children's educational attainment. The following summary draws heavily from the Haveman and Wolfe review.

1. A child's educational attainment is positively related to parent's educational attainment. Another variable describing parental characteristics most commonly used in studies of a child's educational attainment is the educational attainment of the child's parents. Haveman and Wolfe find that "in virtually every study...(parental human capital) is statistically significant and quantifiably important." They further find that a child's educational attainment is more closely related to the educational level of the mother than the father. "Parental completion of high school and one or two years of post-secondary schooling are typically found to have a larger effect on children's schooling than years of parental schooling beyond that level." All of the studies cited by Haveman and Wolfe find that parental education is positively related to their children's educational attainment. Specific studies cited include Hill and Duncan (1987), Krein and Beller (1988), and Case and Katz (1991).

2. Children who grow up in a low-income family tend to have lower educational attainment than children from wealthier families. Haveman and Wolfe find the family income variable, in all studies reviewed except one, to be "positively associated with educational attainment of the child, and the variable is statistically significant in more than half of all cases where a positive relationship is estimated." In their review, Haveman and Wolfe cited a study

by Martha Hill and Greg Duncan (1987) as "one of the most careful explorations of the relationship between family income and children's education." In this study, it was determined that a 10 percent increase in family income (holding all other variables constant) was associated with an increase in educational attainment of less than one percent. However, also cited were Becker and Tomes (1986) who find elasticities in the .01-.02 range. Krein and Beller (1988) and Graham, Beller, and Hernandez (1994) find elasticities of .01-.04.

Other studies reviewed by Haveman and Wolfe which used family income as an explanatory variable on a child's educational attainment include Behrman et al. (1994), Duncan and Hoffman (1990) and Brooks-Gunne, Jeanne et al. (1993). Each of these studies consistently showed family income to be positively related to a child's educational attainment.

3. Growing up in a single-parent family has a negative effect on educational attainment. Haveman and Wolfe cite several studies to support this conclusion. Amato and Keith (1991), in a study on the effects of a child living in a divorced or intact family, conclude that "...on average, having parents who are divorced reduces educational attainment by nearly .2 standard deviations."[1] Haveman and Wolfe also site a study by Lanahan and Sandefur (1994) that "compared the proportion who graduate from high school having grown up in a one-versus two-parent family across five data sets: the differential ranged from 7 to 16 percentage points". Haveman, Wolfe and Spaulding (1991), in a similar study estimated that "the probability of high school graduation of the mean child experiencing two parental separations during ages 6-15 is about five percent lower than that of the child growing up in an intact family." McLanahan and Wojtkiewicz (1992) had similar findings. They report that "a prototypical child living in a one-parent family during ages 14-17 has a 16 percent smaller probability of graduating from high school than a child living in an intact family during these years."

Other studies reviewed by Haveman and Wolfe which included single versus two-

---

[1]This translates into a reduction of about 10 percent in the probability of graduating from high school, and about one-third of a year of schooling attained using the mean and standard deviation of the sample of children included in Haveman, Wolfe, and Spaulding (1991)

parent households are Hill and Duncan (1987), Krein and Beller (1988), Case and Katz (1991), Graham, Beller, and Hernandez (1994), each of which concluded that a child who grew up in a single-parent family attained a lower level of education than a child of a two-parent family.

D. SUMMARY

The demographic and economic characteristics of both academic achievement and educational attainment are well established and thoroughly documented. Important to this thesis is the fact that the same demographic and economic variables which increase an individual's academic achievement also tend to increase educational attainment. An estimation of the number of available "high quality" recruits in a specific geographical area begins with a model which combines these two areas of previous research. This combination is necessary because the military targets those who are "high quality," and at the same time not enrolled in college. Therefore, the model developed in this thesis used to estimate the number of available "high quality" recruits in a specific geographical area will use these well established characteristics.

III. DATA AND METHODOLOGY

A. DATA

This study was conducted using two separate data sets. In stage I of this study, the National Educational Longitudinal Survey of 1988 (NELS 88) was used to obtain parameter estimates (β's) of the relationship between demographic and economic characteristics and the probability the individual is a "high quality," non-college attendee. In stage II of this study, the β's of these characteristics were used to illustrate how an estimate of the number of "high quality", non-college attending individuals at the county level could be obtained. This was done using a second data set, namely the Public Use Microdata Sample (PUMS) 5% sample taken from the 1990 decennial census. The PUMS contains similar demographic variables as the NELS 88, but uses individual level data. Also, the PUMS identifies the county of residence for the individuals included in the data files. For this reason, an illustration of how the parameter estimates can be used to estimate the number of available "high quality" individuals can be done at the county level. Three counties were chosen to demonstrate how county level estimates of available "high quality" recruits could be obtained. The three counties chosen for this simulation are Milwaukee County, WI., Denver County, CO., and Jefferson County, AL. These counties were chosen because they are large urban areas and therefore provided large sample sizes. A detailed discussion of these two data sets follows.

1. NELS 88 Data Set

The 1988 NELS is an individual level survey of 1,035 U.S. schools which tested approximately 25,000 eighth grade students during the base year of 1988. The survey was a two-stage, stratified probability sample with schools selected as the first-stage unit and students within the selected schools as the second-stage. In each school, 26 students were randomly selected with the exception of schools that had fewer than 26 students. In the latter case, all eligible students were included in the sample.

To ensure accurate family characteristics, parents were surveyed and they provided information on their educational expectations for their child, financial support for future schooling and parental involvement in school activities. About 94 percent of the students in

the base year had corresponding parental data. Of the parent survey respondents, 80 percent were mothers, 17 percent were fathers, and the remaining three percent were other male or female guardians. The survey requested that the parent who best knew of the child's learning activities complete the survey.

The NELS 88 data set contains student test scores in the areas of mathematics, reading, science and history/government. Only the mathematics and reading tests were used for this study. The reading test consisted of 21 multiple-choice items to measure student interpretation and comprehension, along with five short passages (one paragraph to one half page). The students were given 21 minutes to complete the reading portion of the test. The mathematics portion of the test lasted 30 minutes and contained 40 items requiring students to make quantitative comparisons, answer word problems, interpret diagrams and complete mathematical calculations.

Although the military uses the AFQT as a selection tool to determine "high quality," this study uses the NELS 88 mathematics and reading portions of the test scores to model "high quality." Because there is a high correlation between the AFQT and the tests given in the NELS 88 (they are both achievement type tests), the NELS 88 test scores should serve as an adequate proxy for the AFQT test scores.

Three follow-up surveys have been completed as of the date of this thesis; 1990, 1992, and 1994. This thesis used the base year outcomes to estimate characteristics of "high quality" and used the third follow-up (1994 survey) to determine college status of individuals.

2. PUMS Data Set

The Public-Use Microdata Samples (PUMS) are microdata files containing the full range of population and housing information collected in the 1990 census. The survey includes 500 occupational categories, age break-outs by single years up to 90, wages in dollars up to $140,000 and other detailed demographic information. Because the samples provides individual level data for all persons living in a sampled household, this study was able to break out individual characteristics of household members such as parental educational attainment.

The PUMS are files which contain records for a sample of housing units with

information on the characteristic of each unit as well as information on the people residing in it. Two separate PUMS are available; a 5 percent sample which identifies all States and various subdivisions within them, including most counties with 100,000 or more inhabitants, and a 1 percent sample which identifies all metropolitan territory and most metropolitan areas with 100,000 or more people, and groups of metropolitan areas elsewhere. This study uses the 5 percent sample files. Each file in the PUMS is a stratified sample of the population. In reality it is a sub-sample of the full census sample (15.9%) of all housing units that received census long-form questionnaires. The 1 percent and 5 percent PUMS were independently drawn samples from the sub-sample.

B. THESIS ESTIMATION METHODOLOGY

1. Overview

Estimation of the number of available 17-21 year old available recruits in a given area begins with the premise that there are two criteria which need to be fulfilled in order for an individual to be qualified for military recruitment. The first premise is that the individual has to meet a mental standard. This standard, as set by the armed forces, is met if the individual is a high school graduate and scores in the top 50th percentile of the AFQT. The AFQT is an achievement test which tests a person's mental ability in mathematics and reading. The second premise is that the individual is a non-college attendee. If an individual attends college, he/she is, for all practical purposes, not available for military recruitment.

An important concept that must be accounted for is the inherent simultaneity which exists between the two aforementioned criteria. Specifically, a person who is of "high quality" (i.e., high test scores) is more likely to attend college than a non-"high quality" person. Therefore, a model which uses *only* "high quality" (or test scores) or *only* non-college attendance as a dependent variable would suffer from simultaneity bias. The model developed in this study attempts to account for the simultaneity by developing a dependent variable which combines the two criteria into one dependent variable which will be known as the "target group" variable.

An individual can fall into one of four categories using the two aforementioned criteria. These categories are; 1) high quality, college attendee, 2) low quality, college

attendee, 3) low quality, non-college attendee, and 4) high quality, non-college attendee. The "target group" variable will therefore have a value of one if an individual is in the fourth of these categories (high quality, non-college attendee) and zero if the individual is in any of the other three categories.

This study was essentially set up in two stages. In stage I, a model was developed to predict the probability that an individual was "high quality" *given* that this individual was a non-college attendee. In stage II, the β estimates obtained in stage I were used along with the same explanatory variables used in stage I (X's) to demonstrate how these β estimates could be used to predict the number of available "high quality" individuals in three selected U.S. counties.

2. Stage I: The Target Group Model

a. Development of the Dependent Variable

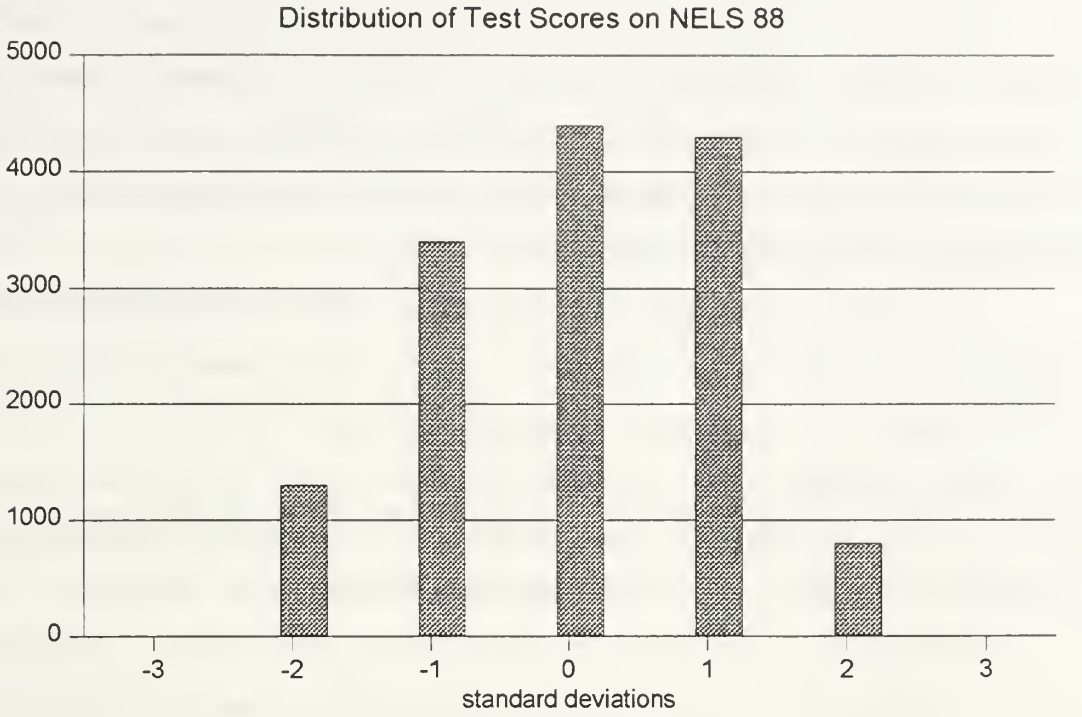
The AFQT measures mathematics and reading abilities and although the AFQT was not used in this study to determine "high quality," it is assumed that the mathematics and reading test scores obtained from the NELS 88 data set serve as adequate proxies for the AFQT.

By combining individual's mathematics and reading test scores obtained from the NELS 88 data, a distribution of test scores was established and standardized with a mean value of zero and a standard deviation of one (Figure 1). Once the mathematics and reading scores were combined and standardized, the binary variable "high quality" was assigned a value of "1" if the individual's test score was in the upper 50th percentile and value of "0" otherwise.

To establish a college attendee/non-college attendee binary variable, the NELS 88 (third follow-up) data set was used (conducted in 1994). With the third follow-up data set, it was possible to establish individual college attendance or non-attendance as the data was collected two years after high school graduation. If a person was in college at the time the data was collected, the variable "college" was assigned a value of one, otherwise the "college" variable was assigned a value of zero.

The "target group" variable used in the model was developed by combining

Figure 1



the two established binary variables "high quality" and "college." The "target group" variable was assigned a value of one if the "high quality" variable was one, *and* the "college" variable was zero, otherwise the "target group" variable was assigned a value of zero. Upon developing the binary "target variable," a probit model was used to estimate the probability an individual will be "non-college" given that the individual is also "high quality." The equation for this probit model is given as:

(3.1) $\mathbf{Pr(T=1) = X\beta + \epsilon}$

where:

T= target group and has a value of "1" if an individual is "high quality" *given* that this individual is also a non-college attendee, and a value of "0" otherwise.

X= various demographic and economic variables (including an intercept)

β= estimation coefficient

ϵ= stochastic error term

The X vector includes well documented demographic and economic characteristics found to be related to academic achievement and educational attainment. Specifically, X includes measures of family income, parental education, family structure, race and gender. Table 2, located in the independent variable section of this chapter, defines all the independent variables included in the X vector.

The decision to use a probit model stems from the fact that the dependent variable ("target group") is binary. If OLS was used to estimate a linear probability model, it would have been possible to have estimated probabilities outside the 0-1 range. In order to bound the estimated probabilities inside the 0-1 interval, the probit model is used. Maximum likelihood estimation of the probit model is produced by interpreting a linear function of the independent variables as an index, in this case potential for being in the "target group." If this potential index exceeds that individual's personal critical value of the index, the person will be in the "target group." Each person's critical value will vary depending on that person's individual characteristics. Some individuals will be very likely to be in the "target group" because they have a high probability of being "high quality" *and* a high probability of being a non-college attendee. Some individuals will be unlikely to be in the "target group" because they either have a high probability of not being "high quality," a high probability of being a college attendee, or both. Critical values, therefore, will be a function of the probabilities of being "high quality" and being a college attendee/non-attendee which is captured in a single dichotomous variable which accounts for both factors, namely the "target group." The critical values of the target group are assumed to be normally distributed which is why a probit model was used instead of a logit model.[2] Table 1 shows the distribution of high quality and college attendance. The percentages given represent the proportion of observations relative to the total number of observations used (10876). Note that, a person who is low quality is much

[2] For further discussion of maximum likelihood techniques see Pindyck and Rubinfeld (1981) chapter 10.

less likely to be a college attendee as 67.6 percent of all low quality individuals were also non-college attendees. Likewise, a person of high quality was much more likely to be a college attendee: 72 percent of all high quality individuals were also college attendees. This illustrates the relationship between high quality and college attendance and underscores the importance of accounting for this relationship in this study.

The explanatory variables used in the probit model are those which have been established in previous studies of individual achievement and college attendance and will be discussed in the following section.

Table 1

Cross Tabulation of Quality and College Attendance Status (number and percent)

	non-college attendee	**college attendee**
low quality	3662 (33.7%)	1750 (16.1%)
high quality	1526 (14.0%)	3938 (36.2%)

source: Author's calculations from NELS 88

b. The Independent Variables

Table 2 lists the definition of the independent variables used in the probit model. The variables were chosen based on previous research conducted in the areas of academic achievement and educational attainment. The frequency distribution of the independent variables can be seen in Table 3.

Table 2
Independent Variables Used in "Target Group" Model

VARIABLE	DEFINITION
MOTHER'S EDUCATION	
Non-High School Mother	=1, mother did not complete high school =0, otherwise
Some College Mother	=1, mother attended college but did not earn a degree =0, otherwise
College Graduate/Advanced Degree Mother	=1, mother earned bachelors, masters or PhD =0, otherwise
FATHER'S EDUCATION	
Non-High School Father	=1, father did not complete high school =0, otherwise
Some College Father	=1, father attended college but did not earn a degree =0, otherwise
College Graduate/Advanced Degree Father	=1, father earned bachelors, masters or PhD =0, otherwise
FAMILY INCOME	
Medium Low Income	=1, total family income is $15,000-$24,999 per year =0, otherwise
Medium Income	=1, total family income is $25,000-$34,999 per year =0, otherwise
Medium High Income	=1, total family income is $35,000-$74,999 per year =0, otherwise
High Income	=1, total family income exceeds $75,000 per year =0, otherwise

Table 2
Independent Variables Used in "Target Group" Model (continued)

RACE	
Black	=1, person is African American =0, otherwise
Hispanic	=1, person is of Hispanic origin =0, otherwise
Asian	=1, person is of Asian origin =0, otherwise
Indian	=1, person is of American Indian origin =0, otherwise
GENDER	
Male	=1, person is male =0, person is female
FAMILY CHARACTERISTIC	
Single Parent Household	=1, family has one parent =0, family has two parents
MISSING/UNKNOWN VALUE DUMMIES	
Mother's Education Unknown	=1, mother's education is unknown =0, otherwise
Mother's Education Missing	=1, mother's education is missing =0, otherwise
Father's Education Unknown	=1, father's education is unknown =0, otherwise
Father's Education Missing	=1, father's education is missing =0, otherwise
Race Missing	=1, race is missing =0, otherwise
Income Unknown	=1, family income is unknown =0, otherwise

Table 3
Proportion of Observations in Each Category From NELS 88

Variable	Proportion with characteristic
Non-High School Mother	14.5%
Some College Mother	18.9%
College Graduate/Advanced Degree Mother	23.2%
Non-High School Father	14.1%
Some College Father	15.9%
College Graduate/Advanced Degree Father	27.8%
Medium Low Income	15.5%
Medium Income	15.5%
Medium High Income	29.7%
High Income	9.1%
Black	12.2%
Hispanic	12.9%
Asian	6.2%
Indian	1.2%
Male	44.2%
Single Parent Household	20.9%

Sample size = 24,599

3. Stage II, Estimation using Stage I Results

In stage II, the β's from stage I were used along with the same explanatory variables used in stage I (X's) and for each observation in the selected counties (using the PUMS 5% data set) a value of the probability of being in the target group was computed. Because the PUMS is set up by households, only those youths who currently reside with their parents will have an associated educational level for a mother and father. For this reason, the age group used to obtain the value of the probability of being in the target group was youths between

the ages of 13 and 18. An assumption is made that those between the ages of 13 and 18 have similar demographic and economic characteristics of those 17-21. The probability of being in the target group was computed using the following equation:

$$(3.2) \quad \mathbf{YBETA}_i = \mathbf{X}_i\beta$$

where **YBETA** is a value between 0 and 1 and is computed for each individual of the three selected counties, **X** is a vector of variables that represent the demographic and economic variables (including an intercept) in the PUMS 5% data file, and β is a vector of coefficients from the NELS 88 model that measures the impact of each of the demographic and economic variables. Because an estimation of the probability that an individual would be in the target group was a dichotomous outcome (an individual was either in the target group or not) it was essential to transform the **YBETA** continuous values into dichotomous values. This was done by comparing the value of **YBETA** to the minus of a random error term drawn from a normal distribution. The following formula was used:

$$(3.3) \quad \mathbf{YHAT}_i = 0 \quad \text{if} \quad \mathbf{YBETA}_i < -\epsilon_i$$

$$= 1 \quad \text{if} \quad \mathbf{YBETA}_i \geq -\epsilon_i$$

If the value of **YBETA** was greater than or equal to minus the intercept threshold value plus the error term, the individual was determined to be of "high quality" *and* a non-college attendee (i.e., in the target group). If this was the case, **YHAT** was set to a value of "1" for that individual. If the value of **YBETA**, on the other hand, was less than the minus of the error term, the individual was determined either to <u>not</u> be of "high quality," be a college attendee or both. In this the case, the **YHAT** was set to "0." A summation of the **YHAT's** for a given county, therefore, represents an estimation of the number of "high quality"/non-college attendee individuals in that county. This estimation, therefore, represents the number of individuals predicted to be in the "target group" for that particular county and represents the size of the available "high quality" recruiting pool for that county.

IV. RESULTS

A. OVERVIEW

In order to estimate the number of "high quality" recruits available in a specific county, this study employs the methodologies from two separate areas of previous research--the determinants of test scores and the determinants of educational attainment. First, estimations of separate models of test scores and educational attainment are produced. The results are consistent with previous studies in these areas. The general results were that the same characteristics which have a strong relationship to student test scores also have the same relationship to educational attainment.

This is not, however, exactly what the military is interested in. The military is interested in the relationship of the characteristics of test scores and *non*-college attendance, as these are the characteristics of the group of people the military is targeting for recruitment. The model developed which accounts for both the relationship between the characteristics of test scores and the characteristics of non-college attendance is what will be known as the "target group" model. In stage I, this "target group" was developed and is specified as:

(4.1) $\mathbf{Pr(T=1) = f(X\beta) + \epsilon}$

where:

T= target group and is assigned a value of "1" if a person is "high quality" and a non-college attendee, and a value of "0" otherwise,

X = demographic and economic variables,

β = coefficient estimators, and

ϵ = stochastic error term.

The results of the "high quality" model and the college model will be discussed as a prelude to the "target group" model to illustrate the effect the demographic and economic variables have on both models. Having established this effect, the results of the "target group" model are then discussed.

B. RESULTS OF THE "HIGH QUALITY" MODEL

The first step is to model the characteristics of a person who is considered to be "high quality." For the purposes of this study, a person is considered to be "'high quality" if that person scored at or above the 50th percentile on the NELS 88 mathematics and reading (combined) tests. A probit model was developed and is given as follows:

(4.2) $\mathbf{Pr(HQ=1) = } f(\mathbf{X}\beta) + \epsilon$

where:

HQ = "1" if a person is "high quality," and a value of "0" otherwise,

X = demographic and economic variables,

β = coefficient estimates, and

ϵ = stochastic error term.

The **X** vector includes the various demographic and economic variables found to be significant in previous studies on academic achievement. The results of the "high quality" model are presented in Table 4. The 'marginal effect' in row 3 compares the probability of being "high quality" for an individual with the specific characteristic with the probability for "base case" person. For this and the following models, the "base case" is a person with the following characteristics: white, female, mother's and father's educational level is completion of high school, family income is below $15,000 and resides in a two-parent household.

As shown, a person with a mother who does not have a high school diploma is 6.96 percentage points *less* likely to be of "high quality" than a person with a mother who has a high school diploma, holding all other variables constant. As a mother's education increases, so too does the probability that a person will be high quality. A person with a mother who attended some college and having earned a college/advanced degree is 10.43 points and 13.00 points more likely, respectively, to be of "high quality" than a person with a mother who has only a high school diploma, holding all other variables constant. A similar relationship exists between quality and the father's education. As the educational level of the father increases from non-high school to college graduate/advanced degree, so too does the probability of that person being "high quality," holding all other variables constant.

Table 4
Results of the High Quality Probit Model

Variable	Coefficient	t-Statistic	Marginal Effect
Intercept	-0.3567	8.06	--
Non-High School Mother	-0.1859**	4.33	-0.0696
Some College Mother	0.2786**	7.94	0.1043
College Graduate/Advanced Degree Mother	0.3472**	9.02	0.1300
Non-High School Father	-0.2403**	5.51	-0.0900
Some College Father	0.1677**	4.52	0.0628
College Graduate/Advanced Degree Father	0.5065**	13.10	0.1896
Medium Low Income	0.2642**	6.00	0.0982
Medium Income	0.3697**	8.35	0.1384
Medium High Income	0.4912**	11.63	0.1839
High Income	0.7090**	11.82	0.2654
Black	-0.6829**	15.40	-0.2557
Hispanic	-0.3594**	9.18	-0.1345
Asian	0.1941**	3.73	0.0727
Indian	-0.5430**	4.23	-0.2033
Male	-0.1239**	5.22	-0.0464
Single Parent Household	0.0206	0.60	0.0077

** Statistically Significant at the 1% level
Log Likelihood Ratio = 15219.513
Concordance Ratio = 74.9%
N = 12,959

Family income is also found to be directly related to the probability of being "high quality." As family income is increased from low to high, the probability of being "high quality" also increases, holding all other variables constant. This is consistent with previous studies of academic achievement.

The race category variables also produce results consistent with previous studies. As seen in Table 4, a person whose race is black, Hispanic and Indian has a lower probability of being "high quality" as compared to a person whose race is white holding all other variables constant by 25.57, 13.45, and 20.33 points, respectively. A person of Asian background is 7.27 points more likely to be considered "high quality" holding all other variables constant. Males were found to be less likely to be "high quality" than females, holding all other variables constant. Although it was found that males were slightly *more* likely than females to be in the upper 50th percentile on the mathematics portion of the test used in the NELS 88 survey, they were also *less* likely to score in the upper 50th percentile in the reading portion. When the mathematics and reading scores are combined, males were found to be 4.64 percent less likely than females to be "high quality", holding all other variables constant. This is consistent with previous research that find males score higher than females on the mathematics portion of achievement tests, but lower than females on the reading portion of achievement tests.

A person who resides in a single parent household is found to be 0.7 points more likely than a person in a two-parent household to be "high quality" holding all other variables constant. However, of all the variables discussed previously, this was the only one not statistically significant. This lack of statistical significance also is consistent with previous studies.

C. RESULTS OF THE COLLEGE MODEL

The next step is to model the probability of a person being a college attendee. Once again, a probit model was developed and is given as follows:

(4.3) $\mathbf{Pr(COLLEGE = 1) = } f(\mathbf{X}\beta) + \epsilon$

where:

COLLEGE = "1" if a person was attending college as determined by the NELS 88 third follow-up data set, and a value of "0" otherwise,

$\mathbf{X}$ = demographic and economic variables,

β = coefficient estimators, and

ϵ = stochastic error term.

It is important to note that the $\mathbf{X}$ vector includes the same demographic and economic variables used in the "high quality" model. The results of the college model are presented in Table 5. The base case is the same for this model as in the "high quality" model.

The results found in this model are consistent with previous studies on educational attainment. An important thing to note is that the same characteristics that increase the probability of a person being "high quality" *also* increase the probability of a person attending college. As parent's educational level and family income increase, the probability of a person being a college attendee also increases, holding all other variables constant. The race variables also have results similar to those in the "high quality" model. In both models, compared to a white person, blacks, Hispanics and those of Indian heritage all had a lower probability of belonging to the specific group (either "high quality or college attendee), while a person of Asian heritage had a higher probability of belonging to each group.

There is one interesting inconsistent result between the two models. This exception is for a youth residing in a single-parent household. Note that in the model of "high quality" a youth who resides in a single-parent household has a slightly higher probability of being "high quality" compared to a person in a two-parent household, holding all other variables constant. This variable, however, is not found to be statistically significant (In fact, it is the only variable in this model to be statistically insignificant). Compare this to the single-parent household variable in the college attendance model. In the college model, a person residing in a single-parent household is found to have a *lower* probability of attending college compared to a person from a two-parent household, holding all other variables constant. Also note that in the college model this variable is found to be statistically significant. A possible explanation for this difference is that family income and the number of parents in a household is highly correlated. A person who resides in a single parent household may not be a college attendee due to a lower family income associated with a one-parent household. It has already

Table 5
Results of the College Probit Model

Variable	Coefficient	t-Statistic	Marginal Effect
Intercept	-0.5192	12.69	--
Non-High School Mother	-0.3294**	8.03	-0.1148
Some College Mother	0.2136**	6.27	0.0745
College Graduate/Advanced Degree Mother	0.4093**	10.67	0.1427
Non-High School Father	-0.1697**	4.22	-0.0592
Some College Father	0.249**	6.90	0.0868
College Graduate/Advanced Degree Father	0.5075**	13.28	0.1769
Medium Low Income	0.2481**	6.19	0.0865
Medium Income	0.3847**	9.38	0.1341
Medium High Income	0.5602**	14.29	0.1953
High Income	0.9691**	15.63	0.3379
Black	-0.0363	0.92	-0.0127
Hispanic	-0.0153	0.43	-0.0053
Asian	0.6211**	12.29	0.2165
Indian	-0.3794**	3.41	-0.1323
Male	-0.2429**	10.53	-0.0849
Single Parent Household	-0.0757*	2.35	-0.0264

**Statistically significant at the 1% level
* Statistically significant at the 5% level
Log Likelihood Ratio = 16146.214
Concordance Ratio = 75.6%
N = 13,822

been established that higher family income results in a higher probability of attending college, so it follows that a person with a single parent would be less likely to attend college.

Generally, as illustrated by the two models, the same characteristics that increase/decrease an individual's probability of being "high quality" also increase/decrease this individual's probability of attending college. This is important to recognize in an effort to predict the number of available "high quality" individuals in a given geographic area. If the relationship that exists between academic achievement and college attendance is ignored, any attempt to predict this number would be biased.

D. RESULTS OF THE TARGET GROUP MODEL

A comparison of the variables and marginal effects in the "high quality" and college models illustrates the notion of the simultaneity that exists between the two. The same characteristics that increase the probability of being "high quality" also increase the probability of being a college attendee. Because the military wants to target those individuals who are of "high quality" they are, in effect, also targeting those individuals who are more likely to attend college, thereby being unavailable for recruitment. In the "target group" model presented in this study, this simultaneity is accounted for by the definition of the "target group."

Table 6 shows the results of estimating the "target group" probit model. The same "base case" as in the previous two models is used. As presented in this table, there appears to be little difference between a youth whose mother or father has no high school degree or some college compared to a youth whose mother or father have a high school diploma, holding all other variables constant. Note, however, as the educational attainment of both parent's increases to college and beyond, a youth is less likely to be in the target group compared to a person whose parents have only a high school degree, holding all other variables constant.

Family income has an interesting effect on the probability of being in the target group. Note that the probability of a youth being in the target group increases, then decreases, as income increases compared to a youth whose family income is low, holding all other variables constant.

Table 6
Results of the Target Group Probit Model

Variable	Coefficient	t-Statistic	Marginal Effect
Intercept	-1.0071	17.86	
Non-High School Mother	0.0234	0.44	0.0056
Some College Mother	0.0058	0.13	0.0014
College Graduate/Advanced Degree Mother	-0.1455*	2.77	-0.035
Non-High School Father	-0.1388*	2.50	-0.0333
Some College Father	-0.0416	0.85	-0.01
College Graduate/Advanced Degree Father	-0.0929	1.80	-0.0223
Medium Low Income	0.0811	1.47	0.0195
Medium Income	0.1225*	2.18	0.0302
Medium High Income	0.059	1.08	0.0142
High Income	-0.2799**	3.45	-0.0715
Black	-0.5503**	8.54	-0.1322
Hispanic	-0.2081**	4.08	-0.05
Asian	-0.3423**	4.83	-0.0822
Indian	-0.0543	0.36	-0.013
Male	0.1343**	4.31	0.0323
Single Parent Household	0.0697	1.56	0.0167

**Statistically significant at the 1% level
* Statistically significant at the 10% level
Log Likelihood Ratio = 8226.558
Concordant Ratio = 61.1%
N = 10,320

All race categories were found to have a lower probability of being in the target group compared to a white person, holding all other variables constant. Males were found to have

a higher probability of being in the target group compared to females, holding all other variables constant. A person who resides in a single-parent household also has a higher probability of being in the target group compared to a person from a two-parent household, holding all other variables constant.

Because the "target group" dependent variable is based on the joint condition of "high quality" and college attendance, it is useful to compare the marginal effects of all three models to illustrate and discuss the simultaneity which exists. Table 7 compares the marginal effects from all three models and will be referred to in a discussion of the marginal effects of the various categories.

As shown in Table 7, a person whose mother is a non-high school graduate is slightly more likely (.56 point difference) to be in the "target group" than a person whose mother is a high school graduate, holding all other variables constant. This is possibly due to the fact that the two factors that determine target group status ("high quality *and* non-college) are working against each other. Although a person whose mother is a non-high school graduate is 6.96 points *less* likely to be of "high quality" and thereby decreasing the probability of being in the target group, this person is also 11.48 points *less* likely to be a college attendee, thereby increasing the probability of being in the target group.

A youth whose mother has some college has exactly the opposite characteristics. This person is 10.43 points more likely to be "high quality" than a youth whose mother is a high school graduate, therefore increasing the probability of being included in the target group. This person, however is also 7.45 points more likely to attend college, thereby decreasing the probability of being in the target group. The end result is a person who is only slightly more likely to be included in the target group (.14 points) than a person whose mother has a high school degree.

Table 7
Comparison of the Marginal Effects From Tables 3,4 and 5

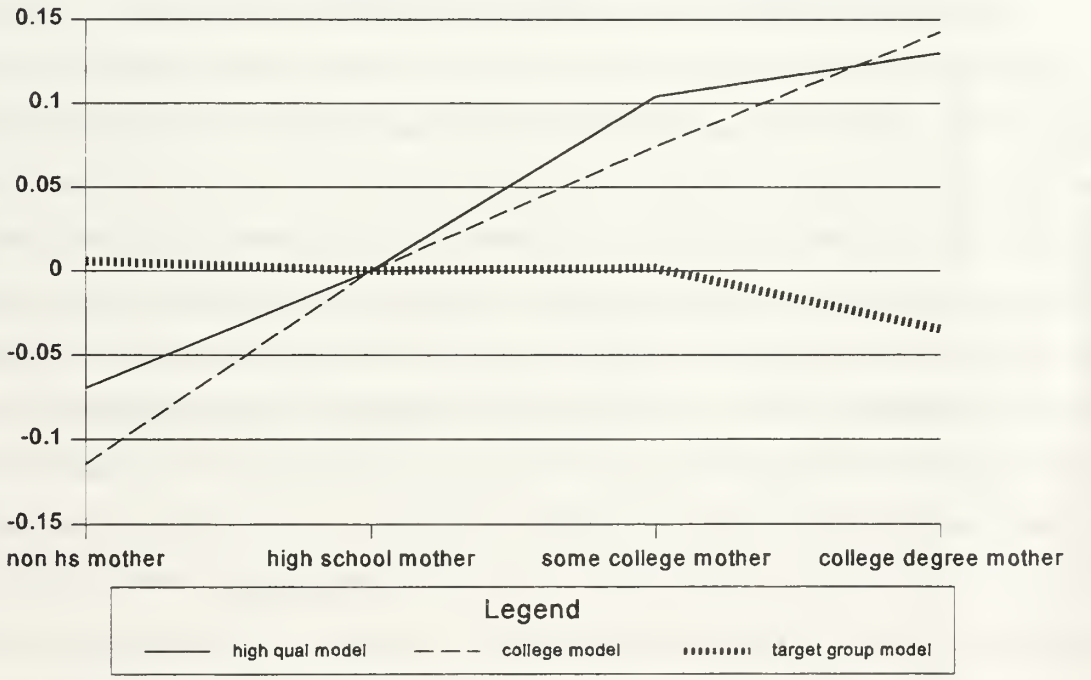
Variable	High Qual Model (Table 3)	College Model (Table 4)	Target Group Model (Table 5)
Non-High School Mother	-0.0696**	-0.1148**	0.0056
Some College Mother	0.1043**	0.0745**	0.0014
College Graduate/Advanced Degree Mother	0.1300**	0.1427**	-0.0350*
Non-High School Father	-0.0900**	-0.0592**	-0.0333*
Some College Father	0.0628**	0.0868**	-0.0100
College Graduate/Advanced Degree Father	0.1896**	0.1769**	-0.0223
Medium Low Income	0.0982**	0.0865**	0.0195
Medium Income	0.1384**	0.1341**	0.0302*
Medium High Income	0.1839**	0.1953**	0.0142
High Income	0.2654**	0.3379**	-0.0715**
Black	-0.2557**	-0.0127	-0.1322**
Hispanic	-0.1345**	-0.0053	-0.0500**
Asian	0.0727*	0.2165**	-0.0822**
Indian	-0.2033**	-0.1323**	-0.0130
Male	-0.0464**	-0.0849**	0.0323**
Single Parent Household	0.0077	-0.0264*	0.0167

** statistically significant at the 1% level
* statistically significant at the 10 % level

As mother's education level rises, it is apparent that the likelihood of being in the target group decreases substantially. A person whose mother has a college or advanced degree is 3.5 points less likely to be in the "target group" than a person whose mother has only a high school degree holding all other variables constant. Again, the effects of both being "high quality" and a college attendee are opposing one another. On one hand, a person whose mother has a college or advanced degree is *more* likely to be in the target group because this person is 13 points more likely to be "high quality" than a person whose mother has only a high school degree, holding all other variables constant. On the other hand, this person is also 14.27 points more likely to be a college attendee, thereby decreasing the probability of being in the target group. In this case, the fact that this person is more likely

Figure 2

Marginal Effects of Mother's Education



to be a college attendee seems to be the dominant influence as the person is less likely to be

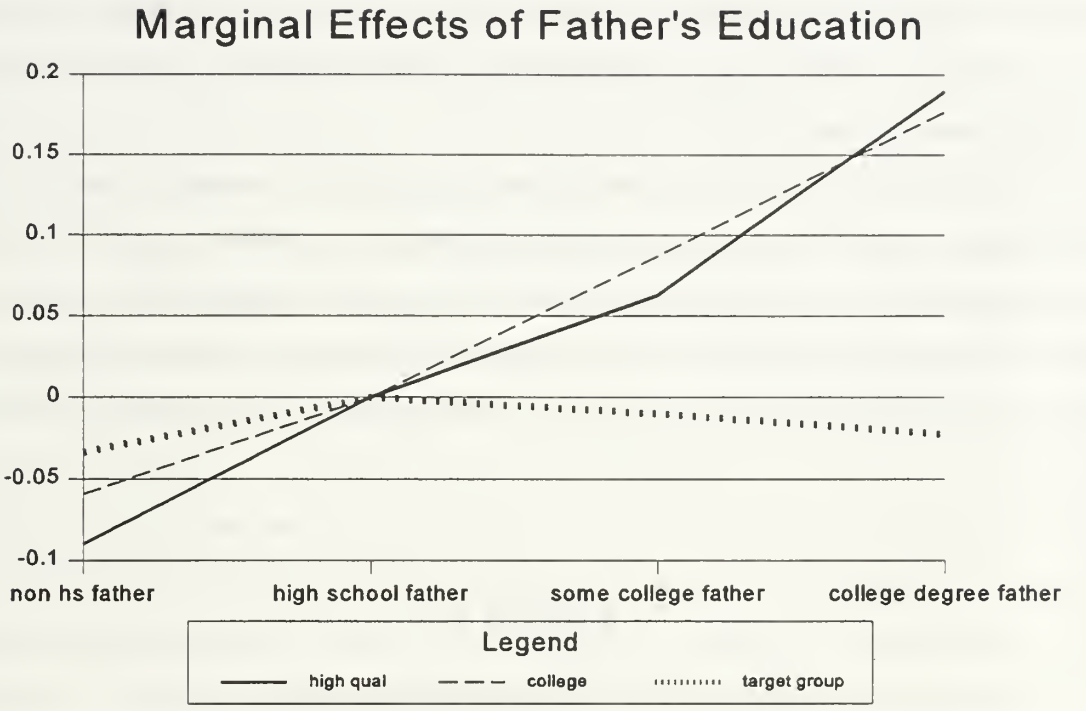
in the target group than a person whose mother is only a high school graduate. Figure 2 illustrates the relationship of the marginal effects of each of the three models. Because the base case for this category variable is a mother having a high school degree, it is assigned a marginal value of "0."

A direct relationship is noted between the educational attainment of the mother and the both probabilities of a youth being "high quality" and a college attendee. The probability of being in the target group, however, shows little difference at lower levels of mother's education. Only when the educational level becomes high does the marginal probability of being in the target group become less than that of a youth whose mother has only a high school degree. Interestingly, the marginal probability of a youth being in the target group becomes lower than that of a youth whose mother has only a high school degree at the same point where the marginal effect of a person who is "high quality" begins to decrease. This further illustrates that at the higher levels of mother's education, the fact that a person has a higher probability of being a college attendee becomes the dominating factor.

The father's educational level has a similar effect on the target group. Both the probability of being "high quality" and the probability of attending college increase as father's education increases. In the target group model, however, a person whose father does not have a high school degree has a lower probability of being in the target group than a person whose father completed high school. Specifically, the non-high school father individual is 3.33 points less likely to be in the target group. The reason for this is that although a person whose father is not a high school graduate is 5.92 points less likely to attend college than a person whose father is a high school graduate, holding all other variables constant, thereby increasing the probability that this person will be in the target group, this person also has a 9 point lower probability of being "high quality" compared to a person whose father is a high school graduate, holding all other variables constant, thereby decreasing the probability that this person will be in the target group.

As father's education increases to some college and college/advanced degree, the probability of this person being in the target group is still less than that of a person whose father has only a high school degree. Specifically, a person whose father's education is some

Figure 3



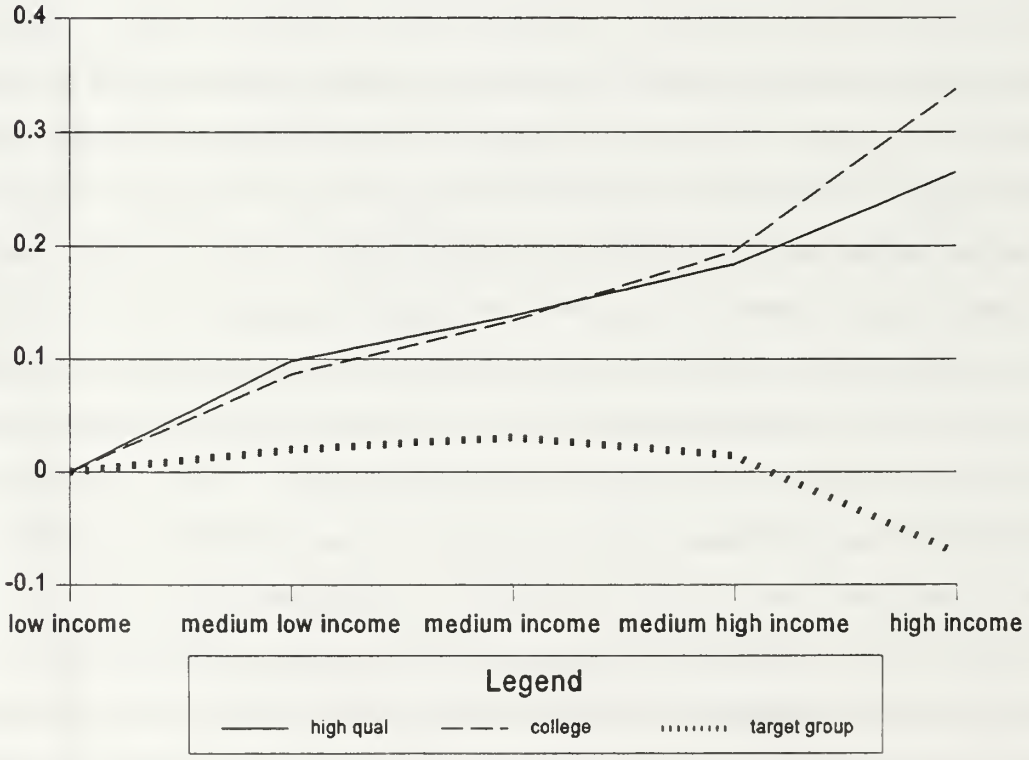
college or college/advanced degree is 1 point and 2.23 points, respectively, less likely to be in the target group than a person whose father has a high school degree, holding all other variables constant. The reason for this decrease in the probability of being in the target group appears to be different than for the non-high school father person. In the case of the non-high school father, this person had a lower probability of being "high quality" that determined the lower probability of being in the target group. In the case of the person whose father had some college or had a college/advanced degree, the determining factor is the fact that the probability of attending college for these individuals is greater, thereby decreasing the probability of this individual being in the target group. Figure 3 illustrates the relationship between the marginal effects of father's education on all three models. As was the case with the marginal effects of the mother's education, there is a direct relationship between the educational attainment of the father and the probability of being "high quality" and a college attendee. Also similar to the mother's education marginal effects is the decrease in probability

of being in the target group as the educational level reaches the college/advanced degree level. A difference between the marginal effects of mother's and father's education is found at the low level of educational attainment. At the lowest level of father's educational attainment, the probability of a person being in the target group is less than that of a person whose father is a high school graduate.

The family income category variable also produced an interesting result. Figure 4 illustrates a comparison of the marginal effect of income on all three models. As family income increases to the medium category, the probability that a person will be in the target group also increases, compared to a person whose family income is in the low category. At the medium high level of family income, the probability of being in the target group begins to

Figure 4

Marginal Effects of Income



decrease. As the income level increases further, the probability of being in the target group decreases substantially. A person whose family income is medium low has a 1.95 point greater probability of being in the target group than a person whose family income is low, holding all other variables constant. Although this person has a 8.65 point greater probability of attending college, thereby decreasing the probability of being in the target group, this person also has a 9.82 point greater probability of being "high quality," thereby increasing the probability of being in the target group. It appears the dominating factor in determining target group status in this case is the "high quality" factor. As family income increases to the medium level, the probability of being in the target group continues to increase compared to the base case individual. Specifically, a person whose family income is at the medium level has a 3.02 point greater probability of being in the target group than a person whose family income level is low, holding all other variables constant. As with the medium low income level, the dominating factor is the fact that this person has a greater probability of being "high quality" (13.84 point difference), thereby increasing the probability of being included in the target group and offsetting the fact that this person is also 13.41 points more likely to be a college attendee than the base case person, which would decrease the probability of being in the target group.

Interestingly, at income levels of medium high and high, the dominating factor influencing target group status changes. At the medium family income level, a person still has a greater probability of being in the target group than the base case person, however, the probability difference decreases from 3.02 points at the medium level to 1.42 points at the medium high level. At these higher levels of income, college attendance status becomes an even more dominating factor. As family income level is increased even further to a high level, the probability of a person being in the target group is less than that of the base case person. Specifically, a person whose family income level is high has a 7.15 point lower probability of being in the target group than a person whose family income is at the low level, holding all other variables constant. Even though a person whose family income level is high has a 26.54 point greater probability of being "high quality than a person whose family income level is

low, thereby increasing the probability of being in the target group, this person also has a 33.79 point greater probability of attending college, thereby decreasing the probability of being in the target group. As Figure 4 illustrates, the marginal effect of being a college attendee begins to increase at a greater rate than marginal effect of being "high quality." As this spread widens, the probability of belonging to the target group decreases substantially, illustrating the fact that at the higher levels of family income the dominating influence of belonging in the target group is the fact that a person has a greater probability of being a college attendee compared to a person whose family income is at the low level.

The race category variable also produced an interesting result. All categories of race included in this study had a lower probability of being in the target group than a person who was white, but apparently for two different reasons. A person whose race was black, Hispanic, or Indian had a 13.22 point, 5 point and 13 point lower probability of being in the target group than a person who was white, respectively, holding all other variables constant. The dominating factor in these cases is the fact that each of these races had a much lower probability of being "high quality," thereby decreasing the probability of being in the target group. A person who was black, Hispanic and Indian had a 25.57, 13.45 and 20.33 point difference (lower) in the probability of being "high quality" than a person who was white, respectively, holding all other variables constant. This lower probability of being "high quality" decreases the probability of a person being in the target group, and although these race groups also had a lower probability of attending college than a person who was white, which would increase the probability of a person being in the target group, the college attendance factor was not enough to overcome the "high quality" factor. The result was a lower probability of being in the target group for all three racial categories compared to a person who was white, holding all other variables constant.

The Asian racial group also had a lower probability of being in the target group than a person who was white; however, the reason is opposite than for the first three racial groups. In the case of the Asian racial group, the fact that this group was substantially more likely to attend college, specifically 21.65 points, was the dominating factor. Although a person who was Asian had a 7.27 point greater probability of being "high quality" which would increase

the probability of being in the target group, the college attendance status was too large resulting in a lower probability of being in the target group.

Males were found to be 3.23 points more likely to be in the target group than females, holding all other variables constant. The dominating factor for this variable was the college attendance factor. Males were 8.49 points less likely than females to attend college, thereby increasing their probability of being in the target group.

The only variable that had the characteristics of an increase in the probability of being "high quality" while, at the same time, also having a lower probability of attending college, was the single parent household variable. As was the case in previous studies of academic achievement, the single parent household variable was not found to be statistically significant in the "high quality" model. However, as in previous studies, this study finds single parent households have a statistically significant effect on educational attainment. Specifically, this study found that a person residing in a single parent household was 2.64 points less likely to attend college than a person from a two-parent household. The fact that the "high quality" factor is not statistically significant, while the college factor is statistically significant, could explain why a person who is from a single parent household was found to be both more likely to be of "high quality," and less likely to attend college compared to a person from a two-parent household, holding all other variables constant.

E. STAGE II RESULTS; COUNTY LEVEL ESTIMATION

As mentioned in chapter III, to illustrate how the β estimates obtained in stage I could be used to estimate the number of available "high quality" individuals at the county level, three counties were selected and a simulation was conducted using the results from stage I. The counties chosen (for illustration purposes only) were Jefferson County, Alabama, Denver County, Colorado, and Milwaukee County, Wisconsin. Table 8 lists the percentages of individuals in each county with the various characteristics and illustrates the complexity of the relationships of the demographic and economic variables. Although Jefferson Co., AL had a substantially greater number of black observations, which would decrease the probability of being in the target group (marginal effect was -13.22 points as compared to whites) it still was predicted to have slightly more target group individuals than the other two counties.

Table 8
Characteristic Proportion for Each County

Characteristic	Jefferson County, AL.	Denver County, CO.	Milwaukee County, WI.
Non-High School Mother	21.6%	31.2%	22.7%
High School Graduate Mother	33.8%	28.0%	34.7%
Some College Mother	23.6%	21.2%	20.4%
College Graduate/Advanced Degree Mother	21.0%	19.6%	22.1%
Non-High School Father	18.8%	31.0%	19.1%
High School Graduate Father	25.3%	22.1%	29.3%
Some College Father	26.2%	16.9%	20.8%
College Graduate/Advanced Degree Father	29.7%	30.0%	30.8%
Low Income	21.3%	25.0%	23.2%
Medium Low Income	14.7%	19.1%	14.2%
Medium Income	15.1%	16.8%	13.3%
Medium High Income	39.9%	30.8%	40.7%
High Income	08.9%	08.3%	08.6%
White	58.8%	52.0%	64.0%
Black	40.2%	20.0%	27.8%
Indian	less than 1%	01.4%	01.1%
Asian	less than 1%	03.7%	02.9%
Hispanic	less than 1%	less than 1%	06.0%
Male	50.7%	54.0%	51.2%
Single Parent Household	29.9%	37.2%	36.0%

Source: Author's calculation from PUMS 5% data set

This variable, however, is not the only variable determining target group status, which means that the combination of all the other variables is great enough to overcome the single black variable resulting in a slightly higher percentage of target group individuals.

Another interesting observation is the fact that Denver Co. CO. has a lower overall educational attainment level for both parents. Denver Co. was found to have a substantially higher percentage of non-high school mother's and father's than the other two counties. Denver Co. was also found to have lower percentages of higher educational attainment for mother's and father's. Mother's with some college was 1.2% more compared to Milwaukee Co., but 2.4% less than Jefferson Co. At the college and advanced degree levels of mother's education, Denver Co. had lower percentages than both of the other two counties. Father's with some college was lower for Denver Co. than the other two counties and about the same for father's with a college or aadvanced degree.

Table 9 shows the results from the simulation which are based on observations at an individual level. The number of observations represents the number of youth between the ages of 13 and 18 in the PUMS files and is of sufficient sample size. The percentages represent the proportion of the youths predicted to be in the target group. Jefferson Co., AL. was estimated to have the largest percentage in the target group, but only by .2 and 1.5 percentage points as compared to Milwaukee Co., WI. and Denver Co. CO., respectively. There isn't a large difference in the percentage of those who were predicted to be in the target group. The homogeneous result makes it difficult to make any concrete statements as to the reason Denver Co. had the lowest percentage estimated to be in the target group, while Jefferson Co. was predicted to have the largest percentage. However, there are some difference in the distribution of characteristics. Denver Co. has more of an "unequal" distribution of characteristics and has the smallest target group, which is what we would expect.

Table 9
County Estimations of Target Population

County	Number of Observations	Number estimated to be in the target group	Percent estimated to be in the target group
Jefferson County, AL.	1,990	190	9.1%
Denver County, CO.	1,084	95	7.6%
Milwaukee County, WI.	1,922	188	8.9%

V. SUMMARY AND RECOMMENDATIONS FOR FURTHER STUDY

A. SUMMARY

This study describes a method of estimating the number of available "high quality" individuals available for recruitment into military service in a local geographical area. Estimating this number is an interesting task complicated by the inter-relationship of test scores and college attendance which exists in the group the military targets. The simultaneity is an important factor which must be accounted for in any estimation of the number of available "high quality" individuals in a geographic area.

The group the military targets is a very unique group. The characteristics that increase an individual's probability of being "high quality" and desirable to the military are the same characteristics that increase an individual's probability of attending college and therefore unavailable for recruitment. The military will probably not be able to recruit individuals at the upper end of the "high quality" spectrum as these individuals are highly likely to be college attendees. However, a market still does exist that satisfies the need for recruiting "high quality" non-college individuals as illustrated in the simulation of the three selected counties.

The general finding of this study is that individuals with very low or very high values of the mother's education, father's education and family income have a lower probability of being in the target group, whereas values of these characteristics near the middle of the distribution correspond to an increase in the probability of belonging to the target group. This was most evident with the family income variable. Estimates, however, are somewhat imprecise due to the inherent multicollinierity between the family background variables.

This study extended past research which estimated the number of available "high quality" recruits by accounting for the simultaneity inherent in the potential recruiting pool. By using this improved method, combined with county level data, recruiting goals which are adjusted for AFQT quality and college attendance status can be updated. Also, a more efficient allocation of recruiting resources across recruiting districts can be achieved.

B. RECOMMENDATIONS FOR FURTHER STUDY

An extension of this study should focus on ways to account for the correlation and interaction of the family background variables. There are a number of ways to approach this problem. One possible approach is a variable reduction technique such a factor analysis.

Another extension to this study would be to estimate the number of available "high quality" for a larger number of counties. This would almost certainly produce a more heterogeneous sample of counties than was found in this study. Also of interest might be an examination of the number of available "high quality" individuals throughout various regions of the United States such as the Northeast, South, Midwest and West.

Another possible follow-on study could explore a redefinition of the target group dichotomous variable. This study combined two binary variables to create the "target group" variable. In this study, a person who scored right at or above the 50th percentile on the mathematics/reading test was included in the target group variable while a person scoring at the 49th percentile was not. Because the "high quality" variable can also be a continuous variable, it would be of value to combine this continuous variable with the binary college attendance variable in a similar model.

An extension of this study could also be done by jointly estimating a test score and college attendance model, explicitly accounting for the simultaneity of both.

LIST OF REFERENCES

Amato, Paul R., and Keith, Bruce, "Parental Divorce and Adult Well-Being: A Meta-Analysis," Journal of Marriage and the Family, Vol. 53, No. 1, pp. 43-58, 1991.

Antel, John, "Mother's Welfare Dependency Effects on Daughter's Early Fertility and Fertility Out-of-Wedlock," Unpublished paper, Department of Economics, U. of Houston, 1988.

Armor, David J., "The Fate of Black America: IV. Why Is Black Education Achievement Rising?," The Public Interest, pp. 65-80, 1992.

Becker, Gary S. and Tomes, Nigel, "Human Capital and the rise and Fall of Families," Political Economy, Vol. 4, No. 3, part 2, pp S1-39, 1986.

Behrendt, Amy, Eisenbach, Jeffery, and Johnson, William R. "Selectivity Bias and the Determinants of SAT Scores," Economics of Education Review, Vol. 5 No. 4, pp. 363-71, 1986.

Behrman, Jere R. *et al.* "A Sequential Model of Educational Investment: How Family Background Affects High School Achievement, College Enrollments, and Choice of College Quality," Department of Economics, Williams College, Williamstown, Ma. 1994.

Bock, R. and Moore, Elsie G.J., "Profile of American Youth: Demographic Influence on ASVAB Test Performances," Office of the Assistant Secretary of Defense (Manpower Installations and Logistics), 1984.

Bradley, R. H., "The Home Inventory: Rationale and Research," in J. Lachenmeyer and M. Gibbs (eds.), Recent Research in Developmental Psychopathology, Gardner, N. Y.: Book Supplement to the Journal of Child Psychology and Psychiatry, pp. 191-201, 1985.

Brooks-Gunne, Jeanne *et al.* "Do Neighborhoods Influence Child and Adolescent Development?," American Journal of Sociology, Vol. 99, No. 2, pp. 353-95, Sept. 1993.

Bumpass, Larry, and McLanahan, Sara, "Unmarried Motherhood: Recent Trends, Composition, and Black-White differences," Demography, Vol. 26, No. 2, pp 279-86, May 1989.

Case, Anne and Katz, Lawrence, "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," National Bureau of Economic Research, Working Paper 3705, 1991.

Cook, Michael D. and Evans, William N., "Families or Schools? Accounting for the

Convergence in White and Black Test Scores," Working Paper, July 1996.

Curtis, E. W., Borack, J. I., and Wax, "Estimating the Youth Population Qualified for Military Service," Navy Personnel Research and Development Center, San Diego, California, 1987.

Desai, Sondale, Chase-Lansdale, P. Lindsay, and Michael, Robert T, "Mother or Market? Effects of Maternal Employment on the Intellectual Ability of 4-Year-Old Children," Demography, Vol. 26 pp. 545-61, 1989.

Duncan, Greg J. and Hoffman, Saul, "Welfare Benefits, Economic Opportunities, and Out-of-Wedlock Births among Black Teenage Girls," Demography, Vol. 27, No. 4, pp. 519-35, Nov. 1990.

Fuchs, Victor R., and Reklis, Diane M., "America's Children: Economic Perspectives and Policy Options," Science, Vol. 255, pp. 41-46, 1992.

Graham, John, Beller, Andrea H., and Hernandez, Pedro, "The Effects of Child Support on Educational Attainment," in Child support and child well being. Eds.: Irwin Garfinkel, Sara McLanahan and Philip Robins. Washington DC: Urban Institute Press, pp. 317-54, 1994.

Hanushek, Eric A., "The Trade-off Between Child Quantity and Quality," Journal of Political Economy, Vol. 100, pp. 84-117, 1992.

Haveman, Robert, and Wolfe, Barbara, "The Determinants of Children's Attainments: A Review of Methods and Findings," Journal of Economic Literature, Vol. 33, pp. 1829-78, 1995.

Haveman, Robert, Wolfe, Barbara, and Spaulding, James, "Childhood Events and Circumstances Influencing High School Completion," Demography, Vol. 28, pp.133-57, 1991.

Hayward, Mark D., Grady, William R. and Billy, John O. G., "The Influence of Socioeconomic Status on Adolescent Pregnancy," Social Science Quarterly, Vol. 73, No. 4, pp. 750-72, 1992.

Hehterington, E. Marvis, Camara, Kathleen, and Featherman, David L., "Cognitive Performances, School Behavior, and Achievement of Children from One-Parent Households," Washington, DC., National Institute of Education, 1981.

Hill, Martha, and Duncan, Greg J., "Parental Family Income and the Socioeconomic Attainment of Children," Social Science Res., Vol. 16, pp. 39-73, 1987.

Hill, Anne M., and O'Neill, June, "Family Endowment and the Achievement of Young Children with Special Reference to the Underclass," unpublished mimeo, 1993.

Hogan, Dennis P., and Kitagawa, Evelyn M., "The Impact of Social Status, Family Structure, and Neighborhood on the Fertility of Black Adolescents," American Journal of Sociology, Vol. 90, No. 4, pp. 825-55, 1985.

Kohn, Melvin L., and Schooler, Carmi, "Work and Personality: An Inquiry Into the Impact of Social Stratification," Ablex Publishing Corporation, 1983.

Koretz, Daniel, "Educational Achievement: Explanations and Implications of Recent Trends," Washington DC.: Congressional Budget Office 1987.

Krein, Sheila F., and Beller, Andrea H., "Educational Attainment of Children from Single-Parent Families: Differences by Exposure, Gender, and Race," Demography, Vol. 25, No. 2, pp.221-34, 1988.

Laurence, M. T., Overby, J. W. II, and Winter, P. E. "The Medical Fitness of American Youth for Military Service," Defense Manpower Data Center, Arlington VA, 1988

Lundberg, Shelly, and Plotnick, Robert, "Effects of State Welfare, Abortion and Family Planning Policies on Premarital Childbearing among White Adolescents," Family Planing Perspectives, Vol. 22, No. 6, pp. 246-51, 275, 1990.

Milne, Ann M., Meyers, Alvin Rosenthal, and Ginsburg, Alan, "Single Parents, Working Mothers, and the Educational Achievement of School Children," Sociology of Education, Vol. 59, pp. 125-39, 1986.

Orvis, B. R., and Gahart, M. T., "Quality Based Analysis Capability for National Youth Surveys: Development, Application, and Implications for Policy," R-3675-FMP, The RAND Corporation, Santa Monica, CA, 1989.

Plotnick , Robert, "The Effects of Attitudes on Teenage Premarital Pregnancy and Its Resolution," American Sociological Review, Vol. 57, No. 6, pp.800-11, 1992.

Simmons, Janet M., Finlay, Belva, and Yang, Alice, "The Adolescent and Young Adult Fact Book," Washington, DC.: Children's Defense Fund, 1991.

Thomas, George W., and Gorman, Linda, "Estimation of High-Quality Military Available and Interested," Department of Administrative Sciences, Naval Postgraduate School, Monterey CA, 1991.

U. S Air Force/Special Study Team, "U.S. Air Force Personnel Force Composition Study: An Analysis of the Effects of Varying Male and Female Force Levels," Washington, DC.: Headquarters, U.S. air Force, March 1985.

Uhlenberg, Peter, and Eggebeen, David, "The Declining Well-Being of American Adolescents," The Public Interest, pp. 25-38, 1986.

Zill, Nicholas, and Rogers, Carolyn C., "Recent Trends in the Well-Being of Children in the United States and Their Implications for Public Policy," in Andrew J. Cherlin (ed.), The Changing American Family and Public Policy, Washington, DC.: The Urban Institute Press, 1988.

Zill, Nicholas, "trends in Family Life and School Performance," paper presented at the Annual Meeting of the American Sociological Association, Pittsburgh, PA, August, 1992.

INITIAL DISTRIBUTION LIST

Number of Copies

1. Defense Technical Information Center.....2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvoir, Virginia 22060-6218

2. Dudley Knox Library.....2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, California 93943-5101

3. Prof. Michael D. Cook.....2
   Code SM/Cm
   Naval Postgraduate School
   Monterey, California 93943-5103

4. Prof. Stephen L. Mehay.....2
   Code SM/mp
   Naval Postgraduate School
   Monterey, California 93943-5193

5. LCDR Timothy J. Duening.....2
   1654 Rangeline Rd.
   Eagle River, Wisconsin 54521